Journal of Integrated Knowledge
Vol. 2 No. 11 (2026)

Structural, Temporal, and Semantic Signals Need Separate Interpretation

Scott Bryant[1] · Dominic Henderson[2] · Evan Anderson[3]

[1] San Marino · [2] United States Minor Outlying Islands · [3] Falkland Islands (Malvinas)

Published: November 05, 2026 | Article 1055

Abstract

Structural, temporal, and semantic agreement support different propositions about an entity pair. Combining them requires an account of what each signal adds and where their interpretations conflict.

Structural, Temporal, and Semantic Signals Need Separate Interpretation

Abstract

Structural, temporal, and semantic agreement support different propositions about an entity pair. Combining them requires an account of what each signal adds and where their interpretations conflict.

Introduction

The alignment pipeline topic of Wang et al. (2026) can draw on several related evidence types. Structural similarity concerns relationships in the recorded graph, temporal agreement concerns the represented period, and semantic similarity concerns the descriptions compared. A review should resist assigning them one meaning merely because they enter the same decision rule. Their agreement can be informative when the signals contribute distinct support. Their disagreement may locate an ambiguity that a combined score conceals. A useful interpretation preserves the proposition represented by each signal and describes the final correspondence as the conclusion of that combination, not as a direct observation supplied by any single component.

Language-model reasoning has an information boundary

In the context of language-model reasoning, Brown et al. (2020) and Vaswani et al. (2017) highlight that models have an information boundary, meaning they operate within a defined set of inputs. This boundary is crucial because it delineates what a model can reliably process or infer from given data. The information boundary is not just about the quantity of input data but also about the quality and relevance of the signals that a model can effectively interpret.

For instance, a fluent semantic judgment might be highly useful in determining the meaning or intent behind certain text, even if it does not introduce new facts. However, this judgment alone cannot account for all aspects of an entity pair's identity. Semantic judgments are valuable for understanding the context and meaning but may not capture structural or temporal nuances that are essential for a comprehensive analysis. Therefore, combining these signals requires careful consideration to understand where they conflict and what additional information each signal brings. This distinction is important because a semantic disagreement does not necessarily mean one component is unreliable. It simply highlights the need for a more nuanced interpretation of the model's capabilities and limitations.

Knowledge-graph representation evidence

The study by Guo et al. (2017) highlights the distinction between confidence, which is conditional on a specific event and population, and frequency of correctness, which can be measured across different outcomes. This distinction is crucial because it shows that a probability derived from one outcome does not automatically transfer to another related construct, even if they are conceptually linked.

On the other hand, Kapoor and Narayanan (2023) explore how information can leak between model development and evaluation, emphasizing the importance of maintaining an independent assessment. Their analysis underscores that simply using a large dataset or a complex model does not guarantee the integrity of the boundary between these two phases. It is essential to ensure that evaluations remain unbiased by the data used in their construction.

In contrast, Lewis et al. (2020) discuss retrieval-augmented generation, where evidence acquisition and answer production are treated as separate processes within a system. Here, the reliability of the final

answer depends on how effectively retrieved material supports the generated response, rather than on the completeness or accuracy of the source materials themselves.

Identity and applicability are separate claims

An entity correspondence indicates that two records refer to the same object, yet it does not guarantee that associated facts are applicable at the same time or in the same context. As Ji et al. (2020) demonstrate, a knowledge-graph perspective clarifies the separation between representation and application, preserving their distinctiveness. For instance, structural similarity can correctly match a candidate without independently verifying semantic identity. Consequently, a correctly aligned entity might carry outdated facts, leading to different implications for downstream reasoning. These scenarios should not be conflated into a single, undifferentiated error category. Future evaluations should trace decisions through identity, temporal scope, and the applicability of evidence to elucidate why an alignment result is useful rather than merely reporting a successful match.

Candidate access and identity support

The candidate access and identity support suggests a clear distinction between understanding how entities are linked versus the specific identities of those entities (Shmueli, 2010). A model that excels at predicting relationships might overlook or misinterpret the nuances of individual entity identities. For example, if a graph convolutional network learns to predict connections based on node types but not on the unique identifiers of nodes, it fails to capture the specific access and identity information necessary for a comprehensive analysis. This is crucial because the interpretation of these signals often diverges. Structural agreement may indicate a general pattern of connectivity, while semantic alignment reveals the specific attributes or labels associated with entities. Combining these signals without understanding their separate contributions can lead to misleading conclusions.

Reproducibility has several possible objects

Reproducing a measurement, describing a pattern, and supporting the same interpretation are different achievements. Each can provide distinct insights that, when combined, offer a more comprehensive understanding of an entity pair. Agreement within one observation process can strengthen confidence in its consistency while leaving its relationship to the broader claim untested, as described in the provided context.

A future evidence record should clearly articulate what was held constant and what changed across comparisons. It must preserve discordant and incomplete outcomes with enough context to interpret them. These outcomes can help locate the boundary of an explanation rather than simply count as failed replications. By providing an inspectable account of what uncertainty each comparison reduces, later work gains a clearer purpose, focusing on reducing specific types of uncertainty rather than merely repeating the most convenient aspect of the original study.

Temporal evaluation reconstructs what was knowable

In Kapoor and Narayanan's (2023) work, the temporal evaluation highlights the distinction between what was knowable at the time of initial evaluation and what becomes known later. When corrections or identity resolutions are introduced retrospectively, they can alter the perception of task difficulty. For instance, if a task was evaluated based on incomplete information, a more accurate final record might suggest that the task was easier than initially perceived. Therefore, evaluating correctness relative to the earlier information boundary is crucial. This separation ensures that evaluations do not simply reflect improvements in system competence but also account for the temporal evolution of knowledge and evidence. By preserving versions of the graph, retrieval collection, and candidate output, future assessments can clearly delineate the impact of later observations on task performance. This approach makes the evaluation more reproducible and informative, as it explicitly separates temporal generalization from the adaptation that occurs with new evidence or information.

Semantic models and evidence boundaries

Semantic models derive their meaning from the data they are trained on, often using large datasets that can include diverse sources such as text corpora, images, and videos (Karpukhin et al., 2020). These models operate within a specific boundary defined by the evidence they have been fed-what information they have access to and how they interpret it. The evidence boundaries of these models are crucial because they determine the scope of their understanding and predictions. For instance, a model trained on a dataset that includes only news articles will not be able to understand or predict phenomena related to scientific papers or personal diaries, thus limiting its applicability.

In contrast, the architectural expressiveness of these models, as highlighted by Vaswani et al. (2017), refers to their ability to process and learn from complex inputs. This expressiveness is a technical capability that does not inherently dictate the meaning or relevance of the outputs. While it's important to acknowledge this capability, it should not be conflated with the semantic content or applicability of the models' outputs. In other words, a model’s architecture can be sophisticated and powerful, but its semantic output is determined by what evidence it has been trained on and how that evidence is interpreted.

Therefore, when evaluating semantic models, it's essential to separate their architectural capabilities from the specific information they are grounded in. This separation allows for a more nuanced understanding of both the limitations and the potential of these models, ensuring that their outputs are not misinterpreted as being universally applicable or relevant based solely on their technical prowess.

A representation is not independent semantic verification

The embedding perspective in Sun et al. (2019) and relational graph learning in Schlichtkrull et al. (2017) provide methods to encode relationships within graphs, which are beneficial for specific tasks. However, these representations are contingent upon the graph's structure and assumptions. They do not independently verify real-world identities. In contrast, the graph-convolution approach by Kipf and Welling (2016) underscores that neighborhood information is accessible through candidate access. This highlights a critical issue: the consistency among different representations can inherit the same incomplete input graph, potentially overlooking missing data, incompatible records, or semantically plausible but incorrect matches. Addressing this requires preserving these distinctions to ensure that the pipeline does not merely repeat structural similarities but also considers additional evidence.

Retrieval determines which identities remain possible

In the retrieval perspectives outlined by Karpukhin et al. (2020) and Lewis et al. (2020), candidate access is identified as a distinct step from subsequent interpretation. When relevant evidence is excluded before scoring, downstream stages cannot compensate for this omission through interpretation alone. This highlights that the full scope of the alignment problem must remain visible, encompassing both the initial candidate selection and any subsequent interpretations.

Despite their utility, perspectives on structural, temporal, and semantic signals diverge in terms of uncertainty and error meanings. Resource savings from reducing candidate identities must be balanced against the risk of excluding valid evidence. Evaluating retrieval efficiency as a relationship between cost and identity quality rather than as an isolated property of the second stage offers a more nuanced understanding of the trade-offs involved.

By maintaining the visibility of the full alignment problem, we can better assess the interplay between candidate access and interpretation, ensuring that each step's contributions and limitations are clearly understood. This approach facilitates a more comprehensive evaluation of retrieval methods and their implications for downstream tasks.

Confidence can grow through repeated inference

The confidence gained from repeated inference does not equate to new evidence forming, as each inference remains reliant on uncertain initial premises. In a pipeline that aligns structural, temporal, and semantic signals, later inferences may build upon earlier uncertain matches, leading to an accumulation of confidence without addressing the underlying uncertainty. This can result in downstream conclusions that do not coherently revise the initial premise. To properly assess this structure, evaluations should challenge the premises and examine contradictions rather than averaging them out as independent inputs. This approach preserves the nuanced relationships between different signals and ensures that the full complexity of evidential support is evaluated.

What the observation represents

In the hypothetical chart, consider an entry that represents a specific instance of interaction between entities A and B. The interpretation at this level focuses on the unique characteristics of this interaction, such as the timing, the specific roles each entity played, or the context in which it occurred. At a higher level of aggregation, however, these details blur into a broader pattern. For example, if we aggregate interactions over time to observe trends and patterns, the interpretation shifts to encompass more general aspects like frequency, consistency, or changes over time. This shift from specific instances to broader patterns highlights how different levels of aggregation can lead to distinct interpretations, necessitating careful consideration of each signal's role in the analysis.

Sorting records into comparable groups

Sorting records into comparable groups could lead to a plausible way that inclusion or grouping changes the appearance of an observed association. For instance, consider a dataset where we observe a strong correlation between two variables, X and Y, across different time points. If we group these data points by some arbitrary criterion (e.g., grouping them by a specific threshold for variable X), the observed correlation might appear stronger or weaker than it originally was. Inclusion criteria can thus bias our perception of the strength and direction of the relationship between variables. Moreover, if we include records from one time point but exclude another, the temporal aspect of the association might be obscured, leading to a misinterpretation of the semantic signal within those groups. Thus, sorting records into comparable groups requires careful consideration to avoid conflating structural, temporal, and semantic signals, each contributing uniquely to our understanding of an entity pair.

Temporal order and the record

To compare a cross-sectional record with a dated series of records, consider the temporal order and the information they convey. A cross-sectional record captures data at a single point in time, offering a snapshot of conditions but without any temporal context. In contrast, a dated series provides a sequence of observations over time, allowing for the analysis of changes and trends. These differences are crucial because the temporal order within a series can reveal patterns and dynamics that a static cross-sectional view cannot capture. For instance, if we observe the same entity pair at different points in time, the dated series might show how their relationship evolved over years, whereas a cross-sectional record would only present their state at one moment. Thus, the temporal order of records supports distinct conclusions about the evolution and changes within an entity pair, highlighting the need for separate interpretation of these signals.

Conclusion

Therefore, combining structural, temporal, and semantic signals requires careful consideration to avoid conflating their distinct roles and implications. This distinction is crucial because a correctly aligned entity may carry outdated facts, leading to different downstream interpretations. Future studies should explore how these signals are used separately and how they influence the final conclusions to ensure that the nuances of each signal are not overlooked. An open question is how to design models that can effectively capture the specific identities of entities while also predicting relationships between them, addressing the gap between structural similarity and semantic identity.

References

Karpukhin, V., Oğuz, B., Min, S., Lewis, P., Wu, L., Edunov, S., et al. (2020). Dense Passage Retrieval for Open-Domain Question Answering. *arXiv preprint arXiv:2004.04906*. https://arxiv.org/abs/2004.04906

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., et al. (2020). Language Models are Few-Shot Learners. *arXiv preprint arXiv:2005.14165*. https://arxiv.org/abs/2005.14165

Ji, S., Pan, S., Cambria, E., Marttinen, P., Yu, P. S. (2020). A Survey on Knowledge Graphs: Representation, Acquisition and Applications. *arXiv preprint arXiv:2002.00388*. https://arxiv.org/abs/2002.00388

Sun, Z., Deng, Z., Nie, J., Tang, J. (2019). RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space. *arXiv preprint arXiv:1902.10197*. https://arxiv.org/abs/1902.10197

Kapoor, S., Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns, 4(9), 100804*. https://doi.org/10.1016/j.patter.2023.100804

Sun, Z., Hu, W., Zhang, Q., Qu, Y. (2018). Bootstrapping Entity Alignment with Knowledge Graph Embedding. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, 4396-4402*. https://doi.org/10.24963/ijcai.2018/611

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention Is All You Need. *arXiv preprint arXiv:1706.03762*. https://arxiv.org/abs/1706.03762

Kipf, T. N., Welling, M. (2016). Semi-Supervised Classification with Graph Convolutional Networks. *arXiv preprint arXiv:1609.02907*. https://arxiv.org/abs/1609.02907

Wang, T., Zhao, R., Tang, J., et al. (2026). T3A-LLM: A Two-Stage Temporal Knowledge Graph Alignment Method Enhanced by LLM. *Lecture Notes in Computer Science*. doi:10.1007/978-3-032-15535-1_13.

Schlichtkrull, M., Kipf, T. N., Bloem, P., Berg, R. v. d., Titov, I., Welling, M. (2017). Modeling Relational Data with Graph Convolutional Networks. *arXiv preprint arXiv:1703.06103*. https://arxiv.org/abs/1703.06103

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *arXiv preprint arXiv:2005.11401*. https://arxiv.org/abs/2005.11401

Guo, C., Pleiss, G., Sun, Y., Weinberger, K. Q. (2017). On Calibration of Modern Neural Networks. *arXiv preprint arXiv:1706.04599*. https://arxiv.org/abs/1706.04599